Some Computer Applications to Problems in Human Genetics*)

DK 681.177 : 576.1

V. A. McKusick

Computer methods have been useful in the following applications in human genetics research (among many): 1) demonstration of what genes are on the same chromosome and the distance separating pairs of loci on the chromosome; 2) demonstration of sex differences in the recombination fraction; 3) assemblage of catalogs of rare recessive phenotypes in man; 4) maintenance of total genealogy of closed populations for such uses as a) determining common ancestor(s) of parents of persons with rare recessive disorders, and b) calculation of coefficients of consanguinity.

EINIGE COMPUTERANWENDUNGEN AUF PROBLEME DER HUMANGENETIK

Der Einsatz von Computern hat sich im Rahmen der humangenetischen Forschung unter anderem auf folgenden Anwendungsgebieten als nützlich erwiesen: 1. Beim Nachweis der auf gleichen Chromosomen liegenden Gene und der Ermittlung der Entfernung der Genloci; 2. beim Nachweis von Geschlechtsdifferenzen in der Rekombinationshäufigkeit; 3. bei der Aufstellung von Katalogen über seltene rezessive Phänotypen beim Menschen; 4. bei der Erstellung von Stammbäumen in geschlossenen Populationen zum Zwecke des Nachweises gemeinsamer Vorfahren von Personen mit seltenen rezessiven Erbkrankheiten und zur Berechnung des Koeffizienten der Blutsverwandtschaft.

Introduction

In the first of these symposia, in 1959, Dr. Talbot and I discussed the problems involved in studying genetic linkage in man. In the third symposium Dr. Murphy and Mrs. Schulze of my group provided more detail on a program for estimating genetic linkage in man. Before embarking on my main topic, let me bring you up to date on studies of linkage which continue at a slow but steady pace.

Genetic linkage studies have as their objective 1) the identification of genetic loci which are on the same chromosome pair and 2) determination of how far apart the loci are on a given chromosome. Mapping the X chromosome and mapping the other chromosomes, the autosomes, have this difference: it is usually clear from the pedigree pattern when a given trait is determined by a gene on the X chromosome. Although it is also clear when a trait is determined by a gene on an autosome, in this latter case it may be any one of 22 pairs of chromosomes that carries the specific locus. In the case of the X chromosome we begin, therefore, with a list of between 50 and 60 traits or diseases determined by separate genes on that chromosome. The question remaining for linkage studies is to map their relative positions. Figure 1 shows the present state of knowledge (1). With the aid of the computer four loci can be positioned as shown. To arrive at this map five linkages were investigated separately. The data have reasonable internal consistency. The linkage of a number of other loci has also been studied, but their positions are still unknown either because the data are thus far too skimpy or because their position is too remote from the marker loci available to permit measurement by the principle of *recombination*, on which linkage mapping is based.

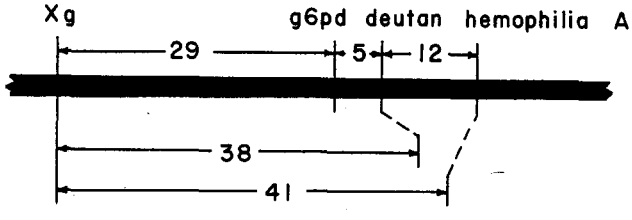


Fig. 1: Relative positions of four genetic loci on the X chromosome of man

*) Paper presented to the 6th IBM Medical Symposium, Poughkeepsie and Brookhaven, N.Y., October 5—9, 1964.

Because of the large number of autosomes and other intrinsic problems, *autosomal mapping* has proceeded much slower. Linkage data on a considerable number of traits have been analyzed by the computer method with negative results; none happened to be located on the same chromosome or at least measurably close on the same chromosome. The computer, however, has demonstrated one new linkage — the Duffy blood group locus and the locus for a form of congenital cataract are on the same chromosome (2). Although suspected by arduous study of this large pedigree, the proof of linkage between these two loci which appear to be rather far apart really required the computer to manipulate the exceedingly complex polynomial which the maximum likelihood method generates from family data of this type. Table 1 lists the pairs of genetic loci which have been shown to be located on the same chromosome together with the distance separating the two loci.

Table 1: Autosomal Linkages in Man

1.	Lutheran blood group system and secretor factor — about 13 map units apart
2.	ABO blood group system and nail-patella syndrome — about 10 map units apart
3.	Rh blood group system and one form of elliptocytosis — about 3 map units apart
4.	Duffy blood group system and one form of congenital cataract — distance not certain
5.	Beta and delta hemoglobin loci — probably contiguous

By computer analysis Renwick and Schulze (3) were able to demonstrate that the recombination fraction (the genetic reflection of chiasma formation) is greater in females than in males — at least for the linkage pair, ABO — nail patella syndrome. Sex differences in recombination fraction had been previously demonstrated in other species, e.g., the mouse, but not in man.

The rest of my discussion will deal with certain applications of computer techniques to population genetics. It is more accurate to say that I will describe some current population studies in which we have needed to enlist the help of computers. The closed populations we have been studying are the several groups of Amish. Before describing the specific investigations now underway, I shall first describe relevant aspects of Amish history and sociology (4).

THE AMISH

Description

Although perhaps not *well* known, the Old Order Amish are *widely* known for their adherence to old-fashioned social and technologic practices. Their closely prescribed manner of dress and use of the horse and buggy are familiar features.

Most Amish are farmers and all are rural-living. Amish society is theocratic. The unit is the church district in which lay clergy regulate all aspects of community life. Religious services are held in the home. The German Bible is used and the south German dialect called Dutch (for *Deutsch)* is spoken within the group. Use of electricity and ownership of modern devices such as the automobile and telephone are forbidden. Resistance to consolidated schools and to education beyond the legal minimum, absolute pacificism, and, in general, separateness from »the world« are other cardinal features. Thus, a common religion and language and a »peculiar people« sense hold the group together.

The Old Order Amish are frequently confused with other varieties of plain people, such as the conservative Mennonites and the Dunkards. There are differences in dress, tonsural practices, etc., but the one feature unique to the Old Order Amish and the one clearest differential feature is holding of religious services in the home. The other groups all have church houses.

From an estimated 8200 in 1905, the Amish population has grown to the present estimated 45,000. In this same period the population of the United States only doubled and part of the increase was due to immigration. Over 80% of Amish live in Pennsylvania, Ohio and Indiana. Over 50% of Amish live in three counties: Lancaster County (Pennsylvania), Holmes County (Ohio) and Lagrange County (Indiana). In Europe the Amish culture was assimilated several decades ago. Except for the group in Ontario, all Amish live in the United States.

History

The Amish sect originated in the Canton of Berne, Switzerland, in 1693 when Jacob Amman led a split from the older and more extensive Mennonite church. Converts were acquired in Alsace, Lorraine, the Palatinate and neighboring areas of southern Germany and eastern France. Many of these converts were Swiss who had moved to these areas in the preceding century. This was a »movement within a movement«.

Migration to eastern Pennsylvania began about 1720 and continued until about 1770. Most present-day Lancaster County Amish are descendants of pre-Revolutionary immigrants, who probably totalled no more than 200 persons. Waves of Amish immigration continued until about 1850, but the later immigrants, finding the land taken up in eastern Pennsylvania, moved on to Ohio and Indiana. The migration patterns account for the peculiarities of distribution of family names and of certain genes in the Old Order Amish of the United States and Canada.

Prospects for Genetic Studies in the Amish

Several features of Amish society are favorable for genetic studies of certain types.

1. It is a defined population, indeed a self-defined population. Although the Amish are often confused with other types of »plain people«, the distinctions are clear to those familiar with the group.
2. It is a closed population. Although some leave, almost no new blood has entered the group since the immigrations.
3. A relatively small number of immigrant ancestors founded each subgroup of the Amish.
4. Genealogic records are excellent. Almost all Lancaster County Amishmen can trace their complete ancestry to immigrants two centuries ago.
5. Undernutrition and infectious disease do not confuse interpretation of findings.
6. Standards of medical care are high. It is especially relevant that diagnostic standards are high.
7. Notable uniformity of socioeconomic circumstances reduces this source of variability.
8. The average level of consanguinity is high.
9. Families are large.
10. Because of their agrarian life, the Amish are immobile. Large kindreds are available for study in a limited geographic area.
11. The Amish are clannish and keep well informed of illness in groups throughout the country through the agency of *The Budget*, a weekly newspaper, and by other means.

Separate Demes

The Amish do not constitute one large genetic isolate. Instead there are a number of more or less separate subisolates, or demes. Deme is the term introduced by Murdock for local endogamous community, or consanguineal kin group. The evidence for separate demes is of several types:

1. History of migration and subsequent separation
2. Family names
3. Collections of rare recessive genes (5)

Family Names as an Indication of Subisolate Formation

Differences in their background, as outlined briefly above, and presumably in the genetic constitution of Amish groups in six areas are reflected in the distribution of family names. (As will be described later, each of four of these six areas is known to have a relatively high frequency of a certain gene which is ordinarily rare.)

In Lancaster County, Pennsylvania, and Holmes County, Ohio, eight names account, in each case, for about 80% of families and no overlap is observed (see Table 2). Here is reflected the pre-Revolutionary and post-Revolutionary origins, respectively, of these groups and the isolation which has been maintained since the immigrations.

Table 2: Old Order Amish Family Names

Lancaster Co., Pa		Holmes Co., O.		Mifflin Co., Pa.	
Stolzfus*	23%	Miller	26%	Yoder	28%
King	12%	Yoder	17%	Peachey	19%
Fisher	12%	Troyer	11%	Hostetler	13%
Beiler	12%	Hershberger	5%	Byler	6%
Lapp	7%	Raber	5%	Zook	6%
Zook	6%	Schlabach	5%	Speicher	5%
Esh**	6%	Weaver	4%	Kanagy	4%
Glick	3%	Mast	4%	Swarey	4%
	81%		77%		85%
Totals: 1106 families, 1957		1611 families, 1960		238 families, 1951	

* Including Stolzfoos ** Including Esch

Mifflin County, Pennsylvania, shows yet another distribution of family names. Although this settlement is also derived from pre-Revolutionary immigrants who settled first in Berks County (Pennsylvania), it was founded by a small group who, for the most part, had names different from those who started the Lancaster County settlement. Only two names, Zook and Beiler (or Byler) appear on the high frequency lists for both Mifflin and Lancaster Counties.

Although family censuses are not available, peculiarities of family name are noted in at least three other areas where discrete immigration in the first half of the last century and relative isolation since then are known to have occurred. These areas and the leading names in each are as follows:

1. Adams and Allen Counties, Indiana: Eicher, Girod, Hilty, Longacher, Neuenschwander, Schmidt, Schwartz, Steury, Wengerd, Wicky.
2. Daviess County, Indiana: Knepp, Stoll, Wagler, Witmer.
3. Perth and Waterloo Counties, Ontario: Albrecht, Jantze, Koepfer, Steckle.

Distribution of Ordinarily Rare Genes as an Indication of Subisolate Formation

Inquiries about various genetic and/or congenital disorders were sent to over 500 physicians practicing in Amish areas of Pennsylvania, Ohio, Indiana, and Ontario. Field trips were made to each of these areas and both physicians and Amish families were visited. Amish informants in communities throughout the country were contacted. Although the methods of these surveys have shortcomings, conclusions about the distribution of certain genes are possible.

Four recessive genetic disorders have been found to have relatively high frequency, each in a different Amish group. These are:

a. *Ellis-van Creveld syndrome* in Lancaster County (Pennsylvania) Amish (6, 7).
b. *Pyruvate kinase deficient hemolytic anemia* in Mifflin County (Pennsylvania) Amish (8).
c. *Hemophilia B (Christmas disease),* an X-linked recessive, in Holmes County (Ohio) Amish.
d. *Limb-girdle muscular dystrophy* in Amish of Adams and Allen Counties, Indiana.

The Usefulness of Inbred Populations for Detection of »New« Recessive Disorders

Relatively speaking, fewer recessively inherited disorders are known in man than in experimental species such as the mouse. In large part this is due to the fact that man is, by and large, an out-breeding animal. In mice if a recessive mutation occurs, it is likely to end up, in 2 or 3 generations, in a homozygote because of close inbreeding. In man the recessive mutation may, by chance, be lost without ever meeting itself in a homozygote in a later generation. Or if a homozygote does occur it may be an isolated case and not find its way into the medical literature, or it may not be clear it is a recessive disorder. Inbred populations provide an increased opportunity for homozygosity of rare recessive genes to occur. A feature of the study of inbred populations particularly attractive to me as a clinician is the opportunity to detect »new« rare recessive diseases. We have, in fact, discovered a »new« disease, a form of dwarfism called *cartilage-hair hypoplasia.* It turns out that cases have occurred in non-Amish persons but the cases were so scattered that no one was impressed that this is a distinct entity and the experience was too limited to prove recessive inheritance.

Where does the computer enter this picture? To answer the questions, just what recessive disorders might you encounter in an inbred population and how are you to know when you have a new one, we have assembled a catalog of rare recessive phenotypes in man. Confining consideration to *rare* phenotypes relieves us of the knotty question of the genetics of conditions such as diabetes mellitus for which a recessive hypothesis has been advanced in the past. Speaking of *recessive* phenotypes and carefully defining the phenotype help avoid the embarrassing situations with those conditions in which some expression is certain or likely in the heterozygote. Maintaining two classes of disorders — those in which recessive inheritance is quite certain and those in which it is possible or likely but not yet proven — permits us to enjoy the heuristic values of a complete enumeration, yet avoid intellectual »sloppiness« of including conditions for which evidence of recessive inheritance is not complete. The catalog has been »computerized« for ease of corrections, additions, deletions, translocations and indexing. The editing and print programs of RICH and OLMER (9) have been very useful and their search program is potentially very useful*. It's a pleasure to acknowledge their help. Over 250 rare phenotypes in man can in our judgment be considered recessive but there must be many more. Studies in inbred populations can uncover as yet undescribed additions to the catalog.

Two main uses to which we are currently putting the computer, in connection with the population genetics of the Amish, are 1) the identification of common ancestor and 2) the estimation of coefficients of consanguinity.

Common Ancestor

Closed populations such as the several Old Order Amish demes are derived from a small number of founders. Each of the founders carried a certain number of rare recessive genes in single dose, that is, in heterozygous state. Each of us carries one or two or more recessive genes, but the probability is that few or none of us carry the same ones. I, for example, may be carrying the PKU gene or the albinism gene. You may be carrying the cystic fibrosis gene or a recessive muscular dystrophy gene. These conditions show up in our children only rarely because when I marry an unrelated person the probability is that her set of bad recessive genes is different from mine. If I marry a cousin, however, the probability of our having the same recessive gene is increased — that probability is 1 in 8, as I will show you — and the probability of my children being homozygous for the gene, getting the bad recessive gene both from its father and its mother, is 1 in 16. Recessive inheritance is, one might say, inheritance from both parents. Dominantly inherited conditions are those which can show themselves when the gene is present in single dose. Recessive conditions are those which show themselves only when the gene is present in double or homozygous state.

When an ordinarily rare recessive disorder is discovered in a closed population with a small number of founders, a logical assumption is that the gene was introduced into the population by a single founder.

Thus far we have detected in the several Amish demes at least a dozen recessive disorders. Which founding father, which immigrant imported the gene? The question resolves itself into, what ancestral couple

*) See also RICH, R. P.: Information handling. (Method. Inform. Med. *4*: 159—163, 1965).

is shared in common by both parents of all sibships with at least one affected member? This is a question the computer is eminently suitable for answering. I can illustrate the logical process by two examples which were worked out arduously by hand.

The Amish deme of Mifflin County (Pennsylvania) contains at least 21 cases of a rare form of anemia due to a *deficiency of the enzyme pyruvate kinase* in the red blood cell. The cases are distributed in 10 separate sibships. The first step in identifying the common ancestor of all 21 parents is to trace the ancestry back to the immigrants, in each case. In order to keep the ancestors straight this is done by setting them out on sheets in which an informative code number is assigned to each — 1 for male ancestor, 2 for female ancestor. Thus, 1 is the father, 2 the mother, 11 the paternal grandfather, 22 the maternal grandmother, etc. Each married couple with affected children is designated by a letter; for example, 1A and 2A refer to the father and mother, respectively, of couple A. Thus, if we find on the sheet a person Jonas Zook with the code 1A2121221 we know immediately that he is the father of the mother of the mother of the father of the mother of the father of the mother of the male member of couple A. In an inbred group such as this, our same Jonas Zook may, of course, appear on the sheets several times. He may for example also appear with the code 1A1121121 (see Fig. 2).

Two further comments: If two persons turn out to be ancestors to all parents of affected children (we have not yet encountered such an example) several possibilities exist: 1) The two ancestors may have been related to each other and therefore carried the same recessive gene. 2) Heterogeneity may exist. Although the condition looks the same, part of the cases may be due to a gene inherited from one ancestor and the cases in other sibships may suffer from a distinct recessive entity inherited from the other ancestor. 3) Only one of the two ancestors may have introduced the gene. If one ancestor is connected to the parents (of affected persons) through appreciably more lines of descent than is the other, he is more likely the person who imported the gene.

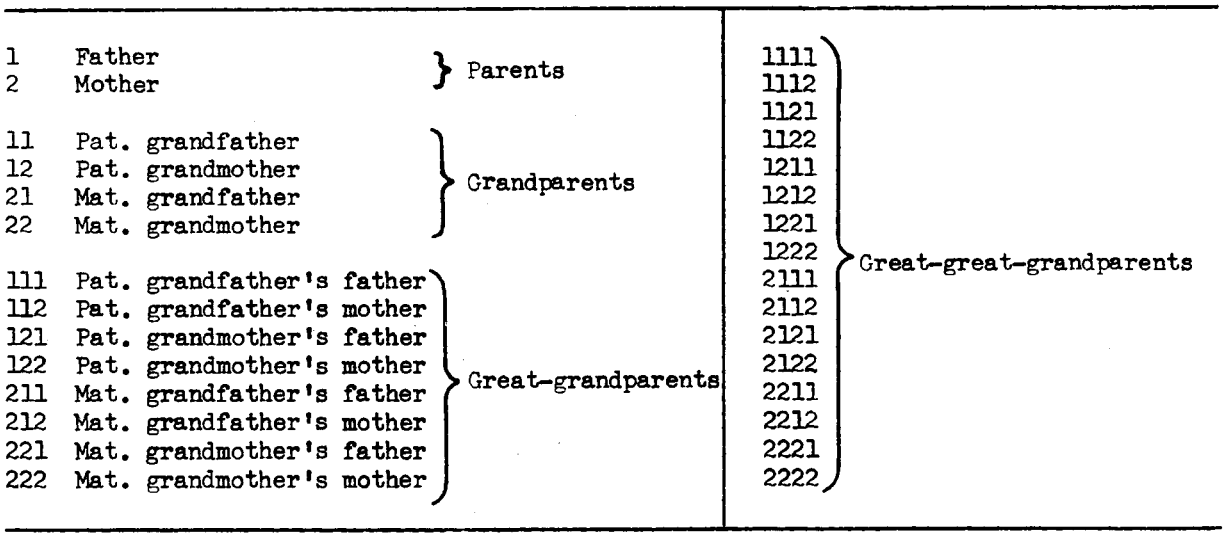


Fig. 2: Code for ancestors

When the schedules are set out for the 21 parents of pyruvate kinase deficiency anemia cases and inspected, five ancestors are found to occur most often, but only one, Strong Jacob Yoder (and, of course, his wife), is ancestral to all 21 parents. The conclusion is that Strong Jacob or his wife carried the gene for this type of anemia.

A more complex example is that of the *Ellis-van Creveld syndrome* in the Amish deme of Lancester County (Pennsylvania). This disorder consists of dwarfism and extra fingers and in about half of cases malformation of the heart. Over 50 cases distributed in 30 sibships have been detected. What founder ancestor is shared in common by all 60 parents? A man named Christian Fisher born in 1757 is an ancestor of 59 of the 60 parents. A man named Nikolas Stoltzfus, who immigrated in 1767, is an ancestor of 57 of the 60 parents. But only Samuel King is an ancestor of all 60 parents. The conclusion is that Samuel King happened to carry the gene for the rare Ellis-van Creveld gene. The pedigree chart (see Fig. 3) tracing 26 of the 30 couples back to Samuel King is a complicated one, looking like a wiring diagram or a map of the London underground.

Consanguinity

For reasons that will be apparent, one wishes to know how closely related married couples are in closed populations, such as these Amish demes. Relatedness is measured by the *coefficient of relationship* (10). Given a married couple (A and B), how closely related are they? Genetically speaking, what is the probability that in person A one of the pair of genes at a given locus is identical by descent (that is to say, was inherited from

the same ancestor) as one of the genes at the same locus in person B? Or, stated more generally, on the average in couples with the same relatedness as couple AB, what proportion of their genes are identical by descent?

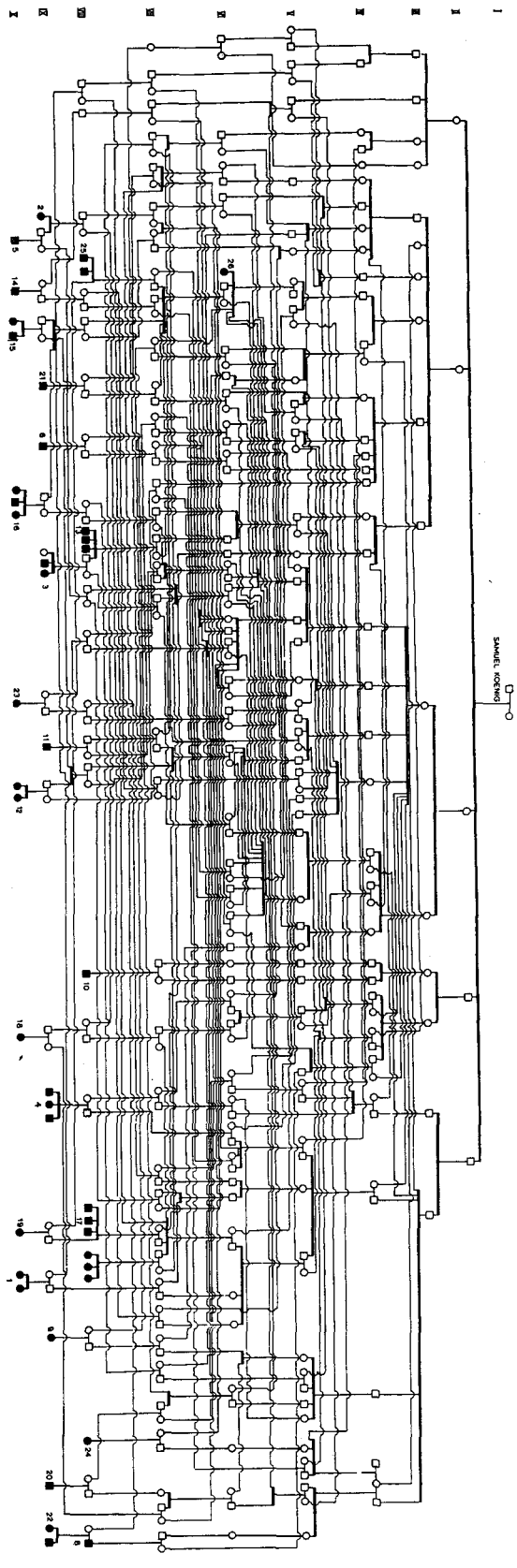


Fig. 3: The descent of both parents of Ellis-van Creveld cases from Samuel King and his wife is demonstrated. The paternity of affected person 26 is in question but it is also uncertain that this infant was in part affected. The affected sibships are numbered. One affected member of sibship 3 is not indicated in this diagram (from McKusick et al.: Bull. Johns Hopkins Hosp. *115*: 306—336, 1964).

The *coefficient of consanguinity* (10) is the measure of inbreeding of an offspring. It is the answer to the question, in a given individual — call him C — what is the probability that the two genes at a given locus are identical by descent (that is to say, are inherited from the same ancestor, having been inherited one from the father and one from the mother)? More generally stated, the question is, in individuals with the degree of inbreeding of person C, on the average, what proportion of genetic loci are homozygous, have pairs of genes identical by descent?

There is a simple relationship between the two values, the coefficient of consanguinity of an individual being 1/2 the coefficient of relationship of his parents.

Half my genes came from my father, half from my mother. The coefficient of relationship of me and my mother is 1/2. One-fourth of my genes came from one of my grandparents; the coefficient of relationship of me and my maternal grandmother is 1/4.

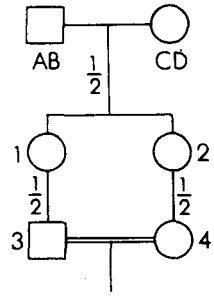


Fig. 4: Calculating the coefficient of relationship (r) of first cousins: $r = {}^1/_8$

In Fig. 4 the estimation of the coefficient of relationship (r) is illustrated for first cousins. The chance that gene A passes to 1 is 1/2, and then to 3 is again 1/2 — in all, 1/4. The chance that 4 gets gene A is also 1/4. The chance that both 3 and 4 get gene A is 1/4 × 1/4, or 1/16. The chance that both 3 and 4 get gene B is also 1/16. These are mutually exclusive possibilities so the final probability, that the same gene (either A or B) is present in first cousins is 1/16 + 1/16, or 1/8. On the average, one-eighth of the genes of first cousins are identical and are derived from the same ancestral source. The same value can be arrived at by the method of path coefficients. One step with a value of 1/2 connects father and daughter and the coefficient of relationship of these persons is 1/2. Connecting uncle and niece there are two paths, each with three steps. The coefficient of relationship $(1/2)^3 + (1/2)^3$ or 1/4. Connecting first cousins there are two paths each with four steps. The coefficient of relationship of first cousins is $(1/2)^4 + (1/2)^4$, or 1/8, the same value we arrived at by another line of reasoning (see Fig. 5).

The coefficient of consanguinity (F) is, as stated earlier, 1/2 the coefficient of relationship. An offspring of first cousin parents has a coefficient of consanguinity of 1/16. This follows because if gene A is given to the offspring by the father, the probability is 1/2 that the mother will also give gene A to the offspring if she has it, and the probability that she has it is 1/8.

In inbred populations two individuals may be related in several different ways. They may be first cousins but also second cousins once removed and third cousins

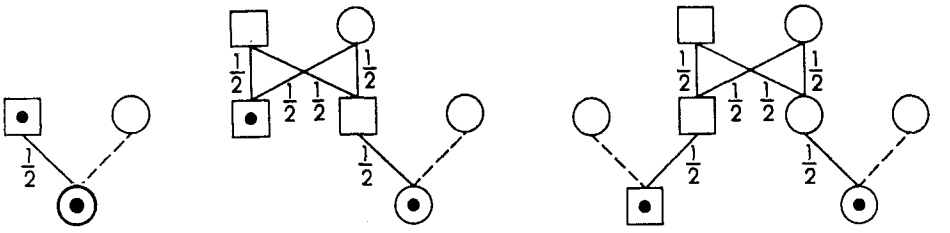
through 2 or 3 or even more other connections. The separate coefficients of relationship are added. Furthermore, the common ancestor may be inbred and this adds slightly to the coefficient of relationship.

Coefficients of relationship are as follows:

First cousins	1/8	.1250
Second cousins	1/32	.03125
Third cousins	1/128	.0078
Fourth cousins	1/512	.00195

two demes being particularly useful for comparison. Estimation of blood group gene frequencies from the phenotype frequencies is facilitated by the computer using the maximum likelihood method which R. A. Fisher introduced for estimates. A program has already been written for this.

Genetic load is a term used to include not only the rare recessive genes which cause grave diseases which prevent reproduction of the homozygote but also so-called lethal equivalents. If there are in a population



Father-daughter
$r = 1/2$

Uncle-niece
$r = (1/2)^3 + (1/2)^3 = 1/4$

First cousins
$r = (1/2)^4 + (1/2)^4 = 1/8$

Fig. 5: The coefficient of relationship (r) of some other close relatives

The mean coefficient of consanguinity is a value which we wish to determine for each Amish deme and for individual church districts. As a guess, the value may lie between second cousins and third cousins — that is to say, it may be the equivalent of all married couples being about second cousins once removed. Calculating coefficients of relationship in inbred populations is a complicated matter because of the complex snarling of the ancestral lines. Dr. Arthur Mange wrote a program for calculating these coefficients and we are adapting his program for use in the Amish. For each individual an identifying code number is assigned and the code numbers of both his parents are indicated in appropriate relationship to his number. Numbers are, of course, assigned to all ancestors as well as the living members of the community.

For both the study of consanguinity and the determination of common ancestors total genealogy is needed. By this I mean complete tracing of ancestry back to the immigrants in the case of each and all members of the community. For the Lancaster County Amish deme this is now 95% complete or more and the data are being assembled on tape for computer manipulation.

Here are other determinations of genetic and demographic significance which can be derived from the computerized total genealogy.

1. To how many immigrant ancestors can the member of each deme be traced and what contribution to the present gene pool of that deme did each make? For example, what proportion of the genes of Lancaster County deme came from Nikolas Stoltzfus? Since he gave his surname to 25% of this group, probably a high proportion of the genes were derived from him also.
2. Average age at marriage.
3. Proportion of persons surviving at 20 who are married at 40.
4. Average number of children.
5. Frequency of twinning and the proportion of like-sex and unlike-sex twins.
6. The sex ratio.

Multiple blood groups are being determined on a random sample of the Lancaster County Amish and similar studies are projected in the Holmes County Amish, these

five genes, each of which in homozygous state reduces the fitness (defined in terms of number of offspring) by 20% as compared with the average, then 1 lethal equivalent is counted in the estimate of genetic load. Genetic load is estimated in inbred populations by comparing the reproductive performance of more inbred persons with less inbred persons. Precocious deaths, including deaths in utero (abortions, stillbirths) and postnatal deaths before age 20 years, provide key data for estimates of genetic load. The number of children surviving to 20 years is another part of the data. We are only beginning to explore the suitability of the Amish group for studies of this type. If we do embark on a full-dress study, the computer will be an essential tool.

Finally, simply assembling medical information of genetic interest on this population is itself a complex operation which will require assistance of the computer for storage and retrieval.

With one of my graduate students (11), we have made preliminary explorations of similar studies in another, much larger, population which shares many of the characteristics of the Amish. These are the French Canadians of the province of Quebec — now prominently in the news for political reasons.

Now numbering about 5 million, the French Canadians descend from a relatively small number of founding fathers. Most of the immigrants came not as family units but as soldiers, fur traders, and voyagers. Shipments of brides, from orphanages for example, provided the wherewithal for procreation. The founding fathers were drawn from rather scattered areas of northern, western and central France. Immigration was interrupted almost completely in 1760 when the British took over. The rate of growth of the French Canadian population has been the greatest of any well recorded white population. They probably constitute the largest white population of comparable genetic homogeneity in the world. Genealogic records are excellent because of the Church records dating back to about 1660. They have since the immigrations remained a relatively closed population by reason of a common religion, tongue, legal code and culture. Consanguinity rates are high even today and dispensation records provide easy access to estimates of same. In parishes of the diocese of Quebec the frequency of first cousin dispensations varies from 2% to 20% with the average about 5%.

It is likely that if one takes any rare recessive disorder found among the French Canadians and traces back all parents of affected children, one will find that one immigrant ancestor can be identified as the carrier, just as it is possible to do in the Amish. It should be possible to estimate genetic load in this population.

Bibliography

(1) McKusick, V. A.: On the X Chromosome of Man. (Washington, D. C.: Amer. Inst. Biol. Sci., 1964).

(2) Renwick, J. H. and Lawler, S. D.: Possible linkage between a congenital cataract locus and the Duffy blood group locus. Ann. Human. Genet. *27:* 67—84, 1963.

(3) Renwick, J. H. and Schulze, J.: Male and female recombination fractions for the nail-patella syndrome. Ann. Human Genet. *28:* 379—392, 1965.

(4) McKusick, V. A., Hostetler, J. A., and Egeland, J. A.: Genetic studies of the Amish. Background and potentialities. Bull. Johns Hopkins Hosp. *115:* 203—222, 1964.

(5) McKusick, V. A., Hostetler, J. A., Egeland, J. A., and Eldridge, R.: The distribution of certain genes in the Old Order Amish. Cold Spring Harbor Symp. on Quant. Biol. *29:* 99—114, 1964.

(6) McKusick, V. A., Eldridge, R., Hostetler, J. A., and Egeland, J. A.: Dwarfism in the Amish. Trans. Ass. Amer. Phycns *77:* 151—168, 1964.

(7) McKusick, V. A., Egeland, J. A., Eldridge, R., and Krusen, D. E.: Dwarfism in the Amish. I. The Ellis-van Creveld syndrome. Bull. Johns Hopkins Hosp. *115:* 306—336, 1964.

(8) Bowman, H. S., McKusick, V. A., and Dronamraju, K. R.: Pyruvate kinase deficiency hemolytic anemia in an Amish isolate. Amer. J. Human Genet. *17:* 1—8, 1965.

(9) Olmer, J. and Rich, R. P.: A flexible direct-file approach to information retrieval, text edit, search and print, on IBM 1401. Proc. 1963 Fall Joint Computer Conference. (Baltimore: Spartan Books, 1963, Pp. 173—182).

(10) McKusick, V. A.: Human Genetics (Prentice-Hall, Englewood Cliffs, N.J., 1964).

(11) Laberge, C.: A prospectus for genetic studies of the French Canadians, with data on consanguinity and blood groups. Bull. Johns Hopkins Hosp., in press.

Author's address: Victor A. McKusick, M.D., Professor of Medicine, The Johns Hopkins Hospital, Baltimore 5, Md., 21 205.